

IT Real-Time training that work for your career.
PROVIDED TRAINING FOR THOUSANDS OF STUDENTS.
SUBJECT, MATERIAL & VIDEOS

**VIDEOS
PROVIDED**

POWER BI

- a) Power Query options [Queries Properties, Query Properties, and Column Properties, Filter Properties]
- b) Case Study-1

MATERIAL

Trainings:
CLASS ROOM
ONLINE



FAST TRACK
ONE ON ONE
PROJECT TRAINING



Address:

Flat No: 506/B
Nilgiri Block
Aditya Enclave
Mytrivanam Area
Hyderabad.

Website & Blog

www.vinaytechhouse.com
www.msbivinay.blogspot.in

Contact Information

+91 9573168449
040 66638869

MSBI

IS, AS, RS & MDS

POWER BI

SERVER, DESKTOP & DAX



WE'VE WORKED WITH A DIVERSE CUSTOMER BASE. HOW CAN WE HELP YOU?
IT Training, Support and Consulting.

Difference between List and Query?

List contain set of values in a column.

Query contains set of columns.

What is expansion at Query?

It will expand the hidden columns in the query.

What are the options at close level in the Power Query?

Close: No changes being saved

Apply: Save the changes applied

Close and Apply: Save the changes and then close.

What are the brief points about Power Query?

- a) The area where we can get data and perform transformations
[Extract and Transform Area]
- b) It uses **Mashup (M Code)** language code for its operations.
- c) Maximum transformations in graphical format
[right click / double click / drag and drop].
- d) If you feel those are not sufficient, then write your own Mashup code / use third party code such as **R and Python**.
- e) For each **query it will generate a Mashup script** which talks about **reading the data onwards till last change steps**
[You can see in the Advanced Editor Area]

What is M Code? Why do we require?

- a) This is a special language designed by Microsoft for **faster and easy transformations**.
- b) It contains set **of classes and methods** [similar to other languages]
- c) We should know the classes and methods before working with them

Sample script code:

Let

```
<variable 1 / source step> = Operation,  
<variable 2/ step 2> = <variable1> + Operation,  
<variable 3/step 3> = <variable 2 > + Operation  
in < final variable 3 / step 3>
```

Explanation:

- a) Each script starts with "**let**"
- b) Each script ends with "**in**" followed by last step / variable
- c) One variable / step separated to other with "**comma**"
- d) Each variable followed by **previous variable**.

Do I need to learn M code in detail?

No need, but understanding [reading] knowledge is added advantage.

Note:

We never write mashup code.

Can we add steps in between?

Yes, you can add in two ways

- a) Graphically, in the query settings, right click any step → step after / before
- b) Through Code, go to code, place the coding statement between the steps

What is Native Query:

The query generated by Microsoft at the time of retrieval and operations on databases.

To see that query settings, go to a step → right click → native query

What is Query Folding:

Taking possible database steps and constructing a query is called query folding.

Useful to perform possible operations at database level (improves performance by running faster) and remaining steps implementation at power query level.

Difference between Copy, Duplicate and Reference?

Copy is a legacy option to Paste it.

Duplicate is Power Query option, but copy and duplicate work in the way.

They take all steps of the copied object and make changes independent later on.

Reference take only source step from the copied or referenced object.

So, if the referencing object changes, referenced object get those changes.

What are the advantages of manual functions in Power Query?

Functions are reusable objects and used for doing set of operations.

We can call multiple times and implement the process all the times.

Usually functions return values.

Vinay Tech House

Working with POWER Query and various options

What is the major role of Power Query in Power BI?

Extract and Transform of data operations, which is suitable for modeling.

How many ways we go to Power Query Area?

- a) Import / Direct query level → Specifying direct Edit / Transform
- b) After import or Direct query → Query editor → Edit Query / Transform Data

How many important options or panes you find?

- a) **Queries Pane** : Queries properties [**Left hand side**]
- b) **Individual Query Pane**: Individual Query and Individual column properties [**Middle**]
- c) **Query settings pane**: Working with operations implemented [**Right hand side**]
- d) **Filter settings Pane**: Against to each column
- e) **Individual Column Properties**
- f) **Multi Column Properties**
- e) **Menu ribbons** [Home, View, PARAMETERS, Add Column, Transform, Help etc.]

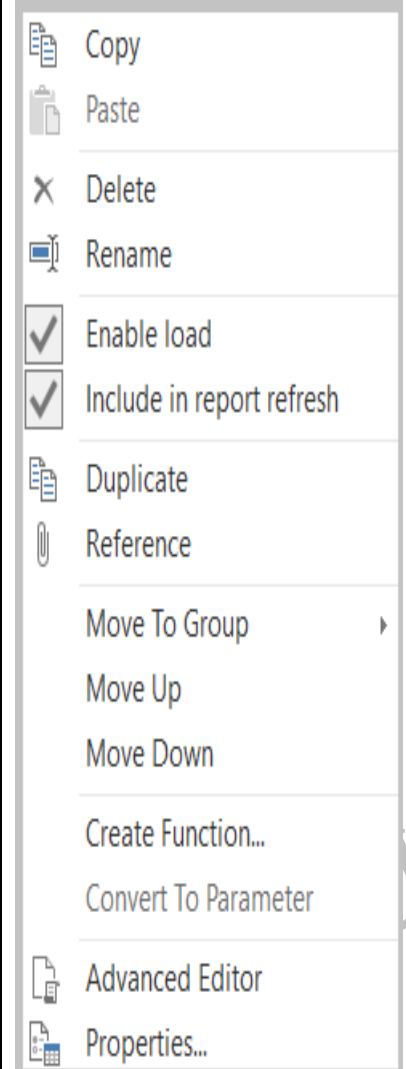
Note: No Modeling ribbon [because this is shaping area]

In which level Power Query Transformations applied?

- a) Queries Level [deal with queries only]
- b) Query Level [deal with Rows Only]
- c) Column Level [deal with individual column values only]
 - Textual column properties
 - Numerical column properties
 - Date column properties
- d) Filter Level [Column cell values filtering]
- e) Multi Column Level [On multiple columns operations]
- f) Menu ribbon options

Vinay Tech House

a) Working on Queries Properties [What are the properties queries section has?]

	<u>Enable Load</u>	a) It will be enabled at report view level b) Displayed at Power BI Dataset level
	<u>Include in Report Refresh</u>	If source changes it will participate in report refresh
	<u>Duplicate</u>	Create duplicate query with all steps Source changes don't affect here
	<u>Reference</u>	Create reference to source and get only source step. Source changes affect here
	<u>Move to group</u>	Create groups for Queries, functions, parameters, lists etc....so that you can work with them easily
	<u>Move up</u>	Moving up the Query / Folder
	<u>Move down</u>	Moving down the Query / Folder
	<u>Create Function</u>	Helps to reuse the function in multiple queries
	<u>Advanced Editor</u>	Opens M-Script for the respective Query

SCENARIO: Duplicate, Copy and Reference practice [Differences between duplicate and Reference?]

Duplicate create an isolated copy with all steps, so changes to the original does not affect here.

Reference create a dependent copy (you can see the query dependencies in Query Dependencies menu item), so changes at original affect reference. You can't drop original without dropping referencing objects.

1. DimStudent query->RC-> Duplicate, rename DimStudent_dup
2. DimStudent Query→ RC→ Reference, DimStudent_ref
3. Convert StudentID to text, Transform Firstname to uppercase in DimStudent Query
4. Observe the Duplicate (no change) and Reference (the changes imposed here) Queries
5. DimStudent query->RC-> delete, it will not allow you because of referencing objects.

SCENARIO: Enable and Include in Report Refresh [What is the impact of Enable and Include Refresh]?

DimStudent→Right Click→ Uncheck Enable Load

DimCourse→Right Click→Uncheck Include in Report Refresh

Close and Apply

Go to Report View, Power BI Dataset and verify the DimStudent →No more visible

Go to DimCourse in Input Excel, add few rows. Go to Power BI Desktop, Click refresh, it will not participate in refresh.

SCENARIO: Move to groups [What is the advantage of groups?]

Arrange the below objects into groups

a) Dimensions

DimCourse, DimLocation, DimStudent, DimCoursemode, DimDate

b) Facts

FactPayment

c) Duplicate and References

DimStudent_dup

DimStudent_ref

Vinay Tech House

SCENARIO: Create function and use it in queries [When do we create and use function?]

Create AnnualDiscountFee function and AnnualDiscount_Inc function and reuse multiple times across queries.

Syntax: **=(argument as datatype, argument as datatype)=> Body of function**

1st Function creation

Queries area (left hand side)→ right click→ New Query→Blank Query→ Specify the below expression.

= (DiscountFee as number)=> DiscountFee * 12

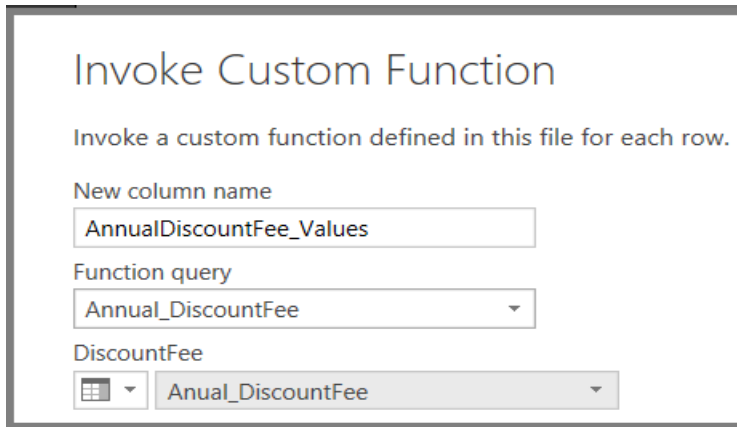
Rename Query as AnnulDiscountFee

2nd Function creation

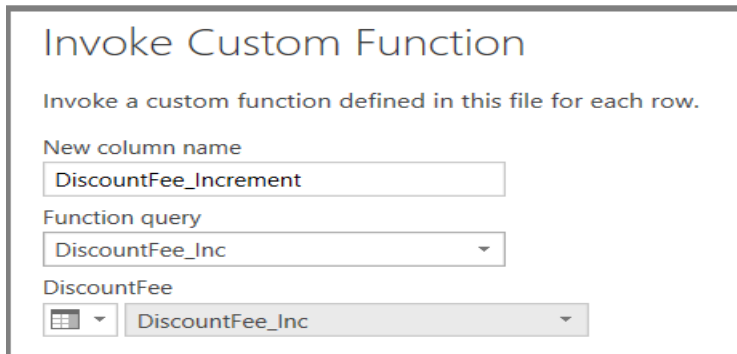
= (DiscountFee as number)=> DiscountFee + DiscountFee * 10/100

Rename query as AnnulDicountFee_Inc

Go to Factpayment query → Click Invoke Custom Function under ADD column menu →



System now adds a new column in the FactPayments and shows AnnualDiscountFee. Similarly invoke custom function AnnualDiscountFee_Inc.



System now adds a new column in the FactPayments and shows AnnualDiscountFee.

b) Working with Query Properties [What are the properties of a query?]

Query data area → Left top drop down and see options

	<p>A Query copy to paste in excel or anywhere</p> <p>Considering the row as header row which has column names</p> <p>Additional column using expression</p> <p>Adding column by entering value or selecting from existing values</p> <p>Invoking functions [Ex: Average of discount amount]</p> <p>Conditional column of single or multiple conditions [if or else if]</p> <p>Add an index column starting with 0 or 1 or required value</p> <p>Choose the columns require in the query for operations</p> <p>Keep the specified start row to end row [to KEEP / REMOVE]</p> <p>Remove duplicates and remove errors.</p> <p>Merge Queries follow join protocol [Inner, Outer and Anti]</p> <p>Append queries follow set theory protocol [Union ALL]</p>
--	---

SCENARIO: Adding Amount to Institute custom column with the below logic

Click Add Custom Column, drag and drop the fields from right hand side queries and create expression [**real time: Additional business columns**]

Custom Column

New column name

Amount_to_Institute

Custom column formula:

= [Discount_Fee]-[Tax amount]

System adds new column and shows the expression result.

Is there any way to implement custom column and conditional column operation?

SCENARIO: Add Conditional Column Rating based on TaxAmount

Click Add Conditional Column, specify like below **[Real-time: Additional business column]**

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name

	Column Name	Operator	Value <small>(i)</small>		Output <small>(i)</small>	
If	<input type="text" value="Tax amount"/>	<input type="text" value="is less than"/>	<input type="text" value="ABC 123"/> <input type="text" value="100"/>	Then	<input type="text" value="ABC 123"/> <input type="text" value="No GST"/>	...
Else If	<input type="text" value="Tax amount"/>	<input type="text" value="is greater than"/>	<input type="text" value="ABC 123"/> <input type="text" value="100"/>	Then	<input type="text" value="ABC 123"/> <input type="text" value="GST Claimed"/>	
Else If	<input type="text" value="Tax amount"/>	<input type="text" value="is greater than"/>	<input type="text" value="ABC 123"/> <input type="text" value="1000"/>	Then	<input type="text" value="ABC 123"/> <input type="text" value="Multiple GST"/>	

Otherwise (i)

What is the difference between custom column and conditional column?

Custom column throws result based on expression, whereas conditional column throws result based on condition satisfied.

SCENARIO: Add Index Columns to the Query [How do we add index columns and impact of it?]

Add index column: Add a new column to the query and specify values

From 0 → The newly added column has values starting from 0

From 1 → The newly added column has values starting from 1

Custom Index → Takes a starting values (Ex: 20) and increment value(1)

1. Add Index column → From 0
2. Add Index Column → From 1
3. Add Index Column → custom Index → Start value: 20 , Increment : 10

Real-time usage:

a) To generate a sequence of value in a column to recognize easily

b) Surrogate key adding as part of data modelling

SCENARIO: Working on Conditional Column [additional example]

Conditional column → Based on condition it add column with required values.

The below example displays ratings for the existing income column values.

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name
Income_Rating

	Column Name	Operator	Value		Output
If	PARTYINCOME	equals	20000	Then	Poor
Else If	PARTYINCOME	equals	30000	Then	Average
Else If	PARTYINCOME	equals	40000	Then	Good

Add rule

Otherwise
Surplus

What are the differences between List and Query?

List	Query
Single column of values	Single or multiple columns
Does not bother about data types	Displays data types and allow us to change
No additional properties to operate [No removal of duplicates, keeping top rows etc...]	We have many additional properties

Scenario: Getting error data in the field?

Add new column [Discount1], write

=DiscountFee + LocationID

Add conditional column,

Take one correct expression,

Take another expression and choose column Discount1

KEEPING AND REMOVING ROWS IN THE QUERY

Scenario: Use the first 100 records available in the input query.

Keep top rows: 100

Scenario: Use the last 100 records available in the input query.

Keep bottom rows: 100

Scenario: Use the rows between 300 and 500 in the input query.

Keep range of rows: 300 and 500

Scenario: Remove top 100 records in the input file and use it.

Remove top rows: 100

Scenario: Remove rows between 300 and 500 in the input file and use it.

Remove range of rows: 300 and 500

Scenario: Remove bottom 100 records in the input file and use it.

Remove bottom rows: 100

Scenario: Remove 2nd record onwards three records and then keep 2 records

Remove Alternate Rows

Specify the pattern of rows to remove and keep.

First row to remove

Number of rows to remove

Number of rows to keep

Scenario: Remove duplicate rows from Factpayments and use it.

Remove duplicates

Scenario: Generating Error Column

Query→ Add custom column→ specify the below expression

= [Discount Fee]+ [LocID]

Scenario: Remove the error data in the ratio column and use it.

[ratio= sum(field)/count(field)]

Remove errors

Scenario: Keep the errors in the ratio column and use it for error analytical reporting.

Keep errors [**People whose calculations are wrong needed report**]

Implementing Set Theory and Joins in Power BI

In Power BI, there are two topics to implement Set Theory and Joins

- a) Append: Set Theory [Union ALL]
- b) Merge: Joins
- c)

Differences between Append and Merge?

Append [Set theory in Power BI]	Merge [Joins in Power BI]
Rows appending from multiple sources	Column based join and get other query columns
Works like Set theory "Union ALL"	Works like Joins in SQL
Row structure should match	Condition column data type should match
Incase column names do not match; it will add as new columns	We have Inner, Outer and Anti Joins only

SCENARIO: Append Operation implementation

If Structures do not match, does it fail?

A) Follows set theory protocol [specially Union All]

B) Merge same structured rows [**no of columns** and **order of data types** matched]

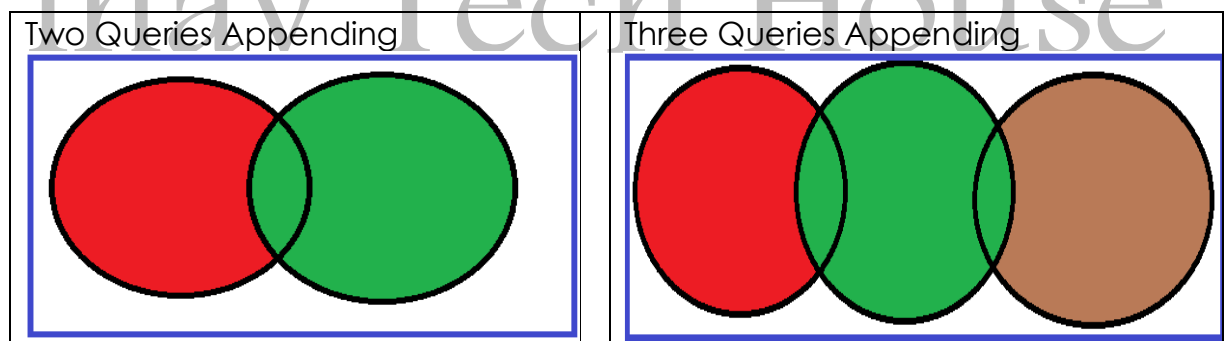
Party: Pid (int), Pnm (varchar(30)), Ploc (varchar(30))

Party1: Pid (int), Pnm (varchar(30)), Ploc (varchar(30))

C) In power BI, these two called as same structured objects

Note: In SQL, there is a possibility of column names mismatch

D) If we specify un structured object, those columns added separately to the result [has nulls for other fields]



Practice:

Take DimCourse, DimCourse_New, DimCourse_upcoming from Institute Business Details

Two queries appending with matching names

Query Editor → DimCourse → Left top corner drop down → Append →

Two tables → Choose DimCourse_New → See the result

Three queries appending with matching and un matching names

Query Editor → DimCourse query → Left top corner drop down → Append →

Three or more tables → choose DimCourse_New and DimCourse_upcoming → Observe

Append

☐ Two tables
 ☒ Three or more tables

Available table(s)

DimCourse (Current)
 DimCourseMode
 DimDate
 DimInstitute
 DimLocation
 DimStudent
 FactPayments
 DimCourses_new
 DimCourses_upcoming

Add >>

Tables to append

DimCourse (Current)
 DimCourses_new
 DimCourses_upcoming

	CourseID	Coursename	Duration	CourseIDs	Coursenames
5	Teradata-N	Teradata Normal Track	40	null	null
6	Teradata-C	Teradata Customized	20	null	null
7	Power BI-F	POWER BI Fast Track	20	null	null
8	Power BI-N	POWER BI Fast Track	30	null	null
9	Power BI-C	Power BI Customized	20	null	null
10	Informatica-F	Informatica Fast Track	30	null	null
11	Informatica-N	Informatica Normal Track	40	null	null
12	Informatica-C	Informatica Customized	20	null	null
13	Azure-F	Azure Fast Track	20	null	null
14	Azure-N	Azure Normal Track	30	null	null
15	Azure-C	Azure Customized	20	null	null
16	Hadoop	Hadoop	30	null	null
17	AI	Artificial Intelligence	50	null	null
18	Data Science	Machine learning	20	null	null
19	DevOps	DevOps	30	null	null
20	null	null		30 Robotics X	Robotics Advanced
21	null	null		50 IOT	Internet Of Things

SQL Joins

1. MULTI TABLE OPERATION and COLUMN WISE OPERATION
2. JOINS UNIQUE FEATURE IS "GETTING MULTIPLE COLS FROM MULTIPLE TABLES"
3. THE RELATIONSHIP ESTABLISHED TO GET COLS FROM OTHER TABLE IS CALLED **JOIN TYPE**.

EID	ENM	DID	DID	DNM
1	VINAY	10	10	IT
2	MADHU	20	30	HR

Source

Joined object [lookup object]

EID	ENM	DID	DID	DNM

Join type	Result	Venn diagram																									
Inner Join [Condition Join]	<table><tr><th>EID</th><th>ENM</th><th>DID</th><th>DID</th><th>DNM</th></tr><tr><td>1</td><td>Vinay</td><td>10</td><td>10</td><td>IT</td></tr></table>	EID	ENM	DID	DID	DNM	1	Vinay	10	10	IT																
EID	ENM	DID	DID	DNM																							
1	Vinay	10	10	IT																							
LEFT Join [Matched+Left Unmatched]	<table><tr><th>EID</th><th>ENM</th><th>DID</th><th>DID</th><th>DNM</th></tr><tr><td>1</td><td>Vinay</td><td>10</td><td>10</td><td>IT</td></tr><tr><td>2</td><td>MADHU</td><td>20</td><td>?</td><td>?</td></tr></table>	EID	ENM	DID	DID	DNM	1	Vinay	10	10	IT	2	MADHU	20	?	?											
EID	ENM	DID	DID	DNM																							
1	Vinay	10	10	IT																							
2	MADHU	20	?	?																							
RIGHT Join [Matched+ Right Unmatched]	<table><tr><th>EID</th><th>ENM</th><th>DID</th><th>DID</th><th>DNM</th></tr><tr><td>1</td><td>Vinay</td><td>10</td><td>10</td><td>IT</td></tr><tr><td>?</td><td>?</td><td>?</td><td>30</td><td>HR</td></tr></table>	EID	ENM	DID	DID	DNM	1	Vinay	10	10	IT	?	?	?	30	HR											
EID	ENM	DID	DID	DNM																							
1	Vinay	10	10	IT																							
?	?	?	30	HR																							
FULL Join [Matched+Left and right Unmatched]	<table><tr><th>EID</th><th>ENM</th><th>DID</th><th>DID</th><th>DNM</th></tr><tr><td>1</td><td>Vinay</td><td>10</td><td>10</td><td>IT</td></tr><tr><td>2</td><td>MADHU</td><td>20</td><td>?</td><td>?</td></tr><tr><td>?</td><td>?</td><td>?</td><td>30</td><td>HR</td></tr></table>	EID	ENM	DID	DID	DNM	1	Vinay	10	10	IT	2	MADHU	20	?	?	?	?	?	30	HR						
EID	ENM	DID	DID	DNM																							
1	Vinay	10	10	IT																							
2	MADHU	20	?	?																							
?	?	?	30	HR																							
CROSS Join [No condition Join]	<table><tr><th>EID</th><th>ENM</th><th>DID</th><th>DID</th><th>DNM</th></tr><tr><td>1</td><td>Vinay</td><td>10</td><td>10</td><td>IT</td></tr><tr><td>1</td><td>VINAY</td><td>10</td><td>30</td><td>HR</td></tr><tr><td>2</td><td>MADHU</td><td>20</td><td>10</td><td>IT</td></tr><tr><td>2</td><td>MADHU</td><td>20</td><td>30</td><td>HR</td></tr></table>	EID	ENM	DID	DID	DNM	1	Vinay	10	10	IT	1	VINAY	10	30	HR	2	MADHU	20	10	IT	2	MADHU	20	30	HR	
EID	ENM	DID	DID	DNM																							
1	Vinay	10	10	IT																							
1	VINAY	10	30	HR																							
2	MADHU	20	10	IT																							
2	MADHU	20	30	HR																							

In case of unmatched data display

Left unmatched = Left join + Where clause right columns evaluated with NULL

Right Unmatched=Right Join +Where clause left columns evaluated with NULL

Power BI Merge Operation implementation

A) Follows Join Protocol [**Column wise Operation and fetch required columns from other object/ query**]

b) It has the below joins

1. **Inner join** → Matched data based on condition
2. **Left join** → Matched data based on condition + unmatched from left
3. **Right join** → Matched data based on condition + unmatched from right
4. **Full join** → Matched data based on condition + unmatched from left + unmatched from right
5. **Left Anti join** → unmatched from left

[In SQL → Left join + where clause right column values evaluated with NULL]

6. **Right Anti join** → unmatched from right

[In SQL → Right join + where clause left column values evaluated with NULL]

What is join? And when do we use it in SQL and other applications?

To bring one or more columns from another table join is required

Establish proper relationships between columns to get desired result.

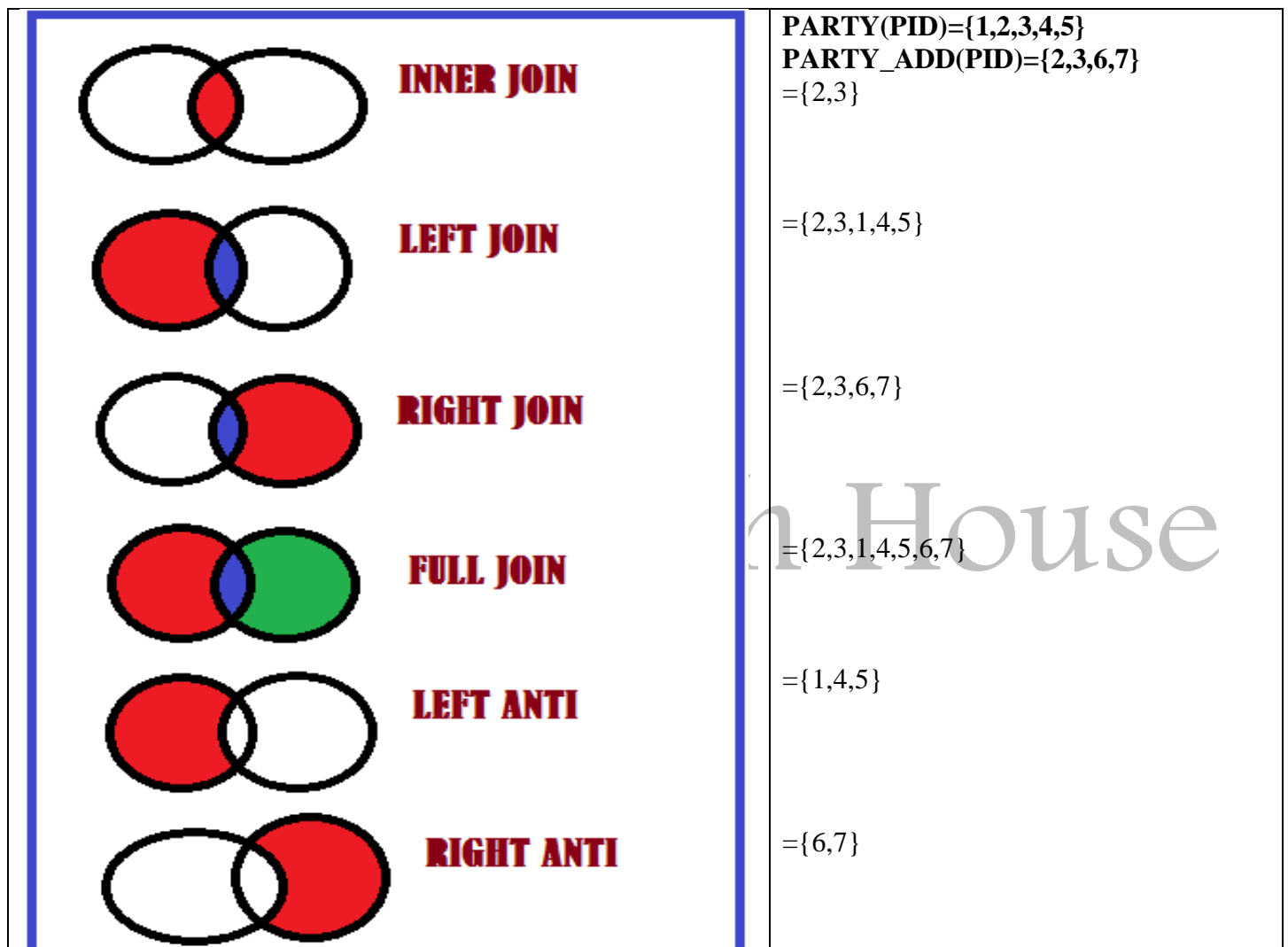
What is the minimal condition required to work with Joins?

Ensure **column data types** matching for joining.

How do we fetch unmatched data in Power BI?

Using anti joins

How do we fetch unmatched data in SQL?



What is the equivalent of ANTI in SQL?

Left Anti=Left join + where clause right columns is NULL

Anti joins required in real-time

- a) For incremental load
- b) To identify left data not available

Inner join:

1. Location Query-> top right corner->Merge Queries-> highlight the LocD column in Location Query
2. Select Geography table and highlight LocID column
3. Select Join (inner) like below

DimLocation

LocID	Locationname	Address
HYD	Hyderabad	506B, Nilgiri Bock,Aditya Enclave, Mytrivanam
VIZAG	Visakhapatnam	602, Sri Nilayam, Near Gayatri College, Vizag
BZA	Vijayawada	Plot no:62, Near PVR, Bandar Road, Vijayawada

+

DimGeography

GeographyID	LocID	Statename	CountryName	ContinentName
H01	HYD	Telangana	India	Asia
V01	VIZAG	Andhra Pradesh	India	Asia
B01	BZA	Andhra Pradesh	India	Asia

Join Kind

Inner (only matching rows)

= Table.NestedJoin("#Changed Type1",{"LocID"},DimGeography,

	AB _C LocID	AB _C Locationname	AB _C Address	DimGeography
1	HYD	Hyderabad	506B, Nilgiri Bock,A...	Table
2	VIZAG	Visakhapatnam	602, Sri Nilayam, Ne...	Table
3	BZA	Vijayawada	Plot no:62, Near PV...	Table

Right Anti join:

1. Location Query-> top right corner->Merge Queries-> highlight the LocD column in Location Query
2. Select Geography table and highlight LocID column
3. Select Join (right anti) like below

DimLocation

LocID	Locationname	Address
HYD	Hyderabad	506B, Nilgiri Bock,Aditya Enclave, Mytrivanam
VIZAG	Visakhapatnam	602, Sri Nilayam, Near Gayatri College, Vizag
BZA	Vijayawada	Plot no:62, Near PVR, Bandar Road, Vijayawada

DimGeography

GeographyID	LocID	Statename	CountryName	ContinentName
H01	HYD	Telangana	India	Asia
V01	VIZAG	Andhra Pradesh	India	Asia
B01	BZA	Andhra Pradesh	India	Asia
W01	Warangal	Telangana	India	Asia

Join Kind

Right Anti (rows only in second)

Additional Examples

1. EMP Query-> top right corner->Merge Queries-> highlight the DID column in EMP Query
2. Select Dept table and highlight DID column
3. Select Join (inner) like below

Merge

Select a table and matching columns to create a merged table.

EMP

EID	ENAME	DID	ELOC
1	VINAY	10	HYD
2	MADHU	40	MUM

DEPT

DEPTID	DNAME
10	IT
20	HR






Join Kind

Inner (only matching rows)

The selection has matched 1 out of the first 2 rows.

OK

Cancel

  f_x		= Table.NestedJoin(dbo_EMP,{"DID"},DEPT,			
	12_3 EID	A^B_C ENAME	12_3 DID	A^B_C ELOC	 DEPT 
1	1	VINAY	10	HYD	Table
2	2	MADHU	40	MUM	Table
3	null	null	null	null	Table

Right click Table in DEPT, you will find three options

- Copy—To create copy of the column/ object
- Drilldown→ To drill and show the respective value
- Add new query→ To create new query

Note: To see the deptname column values, click the right arrow marks at DEPT and choose the column deptname.

Joins extensions in Power BI

Power BI added Fuzzy match for joining operations.

What is Fuzzy match?

Similarity match

Which columns suitable for Fuzzy Match?

Differences between Fuzzy Match and Normal Join?

What are the various fuzzy match options available?

Vinay Tech House

Class room example for fuzzy match?

Use FactPayments CourseID matching with DimCourse table Course Name matching

a) Working with Field Properties [Field Transformations]

There are three types of fields, and the properties slightly different from one to another based on data type.

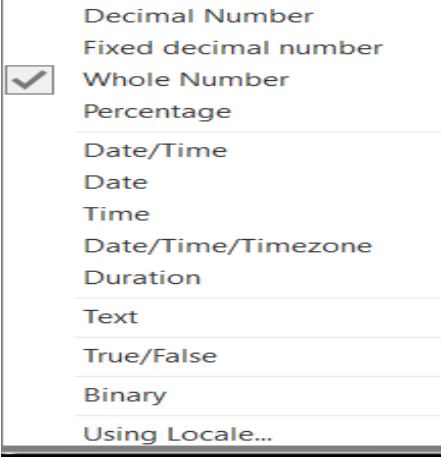
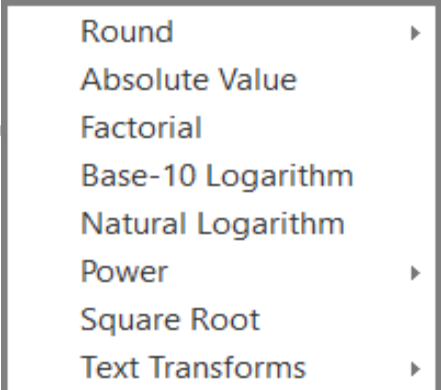
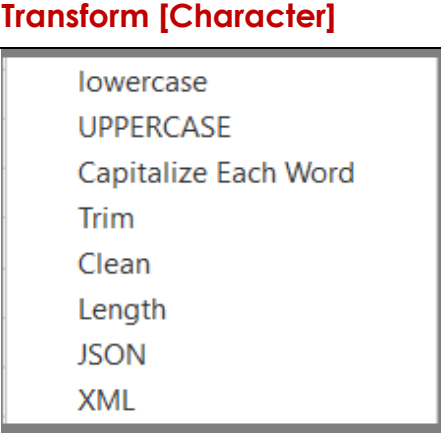
- a) **Numeric column** properties [Decimal, Whole]
- b) **String column** properties [Additional property is split]
- c) **Date column** properties [close to numeric]
- d) **Multi column** properties [additional is merge columns]

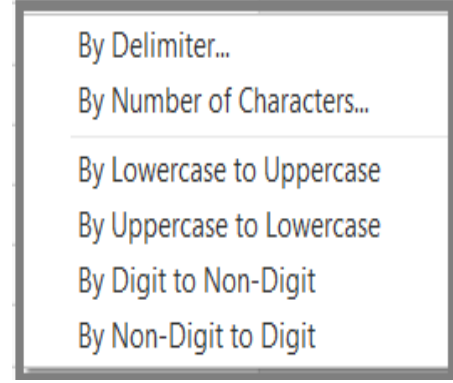
Go to any Column [**Numeric column**] → Right Click Properties →

Copy	Creates copy of the column
Remove	Removes the column
Remove Other Columns	Removes other columns than highlighted
Duplicate Column	Creates a duplicate of column
Add Column From Examples...	You can add from examples given by Microsoft
Remove Duplicates	Remove column duplicate values
Remove Errors	Remove column error values
Change Type	To convert into another data type
Transform	To transform into another format (Ex: uppercase)
Replace Values...	Replaces column values
Replace Errors...	Replace error values
Split Column	Splits the column characters based on value or delimiter
Group By...	Create groups for the cell values
Fill	Fill the values down or up with the last value
Unpivot Columns	Transpose columns to rows
Unpivot Other Columns	Change column name
Unpivot Only Selected Columns	Move to another Query
Rename...	Allow us drill down on the column
Move	The values of column can become a Query
Drill Down	
Add as New Query	

What is the difference between “drilldown” and “add as new query?”

Drilldown will construct a list from the existing (**replaces query with list**), whereas add as new query will **construct a new list**.

<u>Change Type Options</u> 	<u>Importance of data types used for calculations</u> <u>Refer to Data types topic [previous topic]</u>
<u>Transform [Numeric]</u> 	<u>What are the round options available?</u> <u>Differences between Round, Round Up, Round Down?</u> <u>Real-time Observation:</u> <u>What are the Power Options available?</u> Power, Square, and cube.
<u>Transform [Character]</u> 	<u>Differences between Trim and Clean?</u> Trim removes spaces whereas clean removes junk characters <u>What kind of data suitable for JSON?</u> The fields having attribute (key property) and value.
<u>Split data values</u>	

	<p>Based on separator (pipe,comma etc..) splits text.</p> <p>Based on the specified number of characters splits text.</p> <p>abcKLR: abc KLR [Lower to Upper split] KLRabc : KLR abc [Upper to lower]</p> <p>2000ABC: 2000 ABC [digit to non-digit] ABC2000: ABC 20000 [non-digit to digit]</p>
---	--

What are the moving column options available?

Down, up, To Beginning, and To End.

What are the additional options you will find when you work with Strings?

Split, Character Transforms

What are the additional options you will find when you work on Multiple Columns?

Merge

Scenario: Implement the below operations on DimStudent table

- a) Highlight → right click
 1. Duplicate DiscountFee column
 2. Rename it Discount_Fee_new
 3. DiscountFee→Right Click→Change type→ Decimal
 4. Apply [Simply save the steps]
- b) Highlight Firstname→ Right Click
 1. Transform-> Uppercase/ lower case
 2. Apply
- c) Remove the step in the right hand side corner (Duplicate Column) and see the preview.

SCENARIO: Creating list of non-duplicate values in Location

List contain set of values

1. Take copy of Query [Factpayments]
2. Highlight LocationID and remove remaining
3. Remove duplicates
4. Add as new Query.

Note: It will create a list and we can find in the left hand side with list symbol

SCENARIO: Converting list into table

Go to edit queries→ select table→ click TO Table from ribbon→ specify delimiter (in case you have to get multiple columns in the table, now it is NONE (No delimiter))

SCENARIO: Adding new custom column and splitting data

Source:

Add Column→ New Custom Column→ Column name: Course_Mode

Value: = [CourseID] & "," & [ModelID] & "," & "INDIA"

Take a copy of the Query

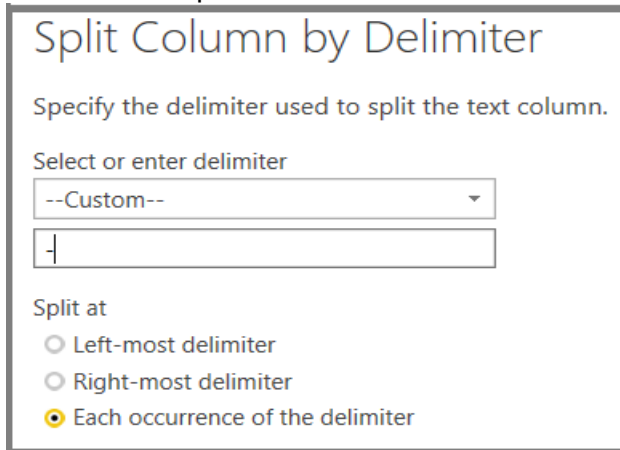
Remove other columns by keeping Course_Mode

Convert Course_Mode into LIST (RC-> Add New Query)

Click TO TABLE and Specify Delimiter (Comma)

SCENARIO: Split based on number of characters

CourseID-> Split-> based on – →to see the coursename and type separately



SCENARIO: Merge columns based on "space"

Use the previous scenario split columns, highlight both columns→Merge→Choose delimiter [this time "space"]

How many types of grouping available in Power BI?

- a) Power Query level grouping: **Automatic value grouping and support aggregations.**
- b) Power BI Dataset level grouping: **Manual value grouping** →List group or Bin group

What is the difference between Trim and Clean?

Trim remove spaces where as clean removes junk characters

How many Remove duplicate or Error options available and how they are different?

Remove duplicates at Query level → Row duplicates removed

Remove duplicates at column level → Column cell value duplicates removed

SCENARIO: Grouping on Institute and finding Count of students, Sum and Average of Discountfee

InstituteID → Right click → Group By and implement the below

Group By

☐ Basic ☒ Advanced

Specify the columns to group by and one or more outputs.

Group by

InstituteID

Add grouping

New column name	Operation	Column
Count_of_Students	Count Rows	
Total_DiscountFee	Sum	Discount_Fee
Average_DiscountFee	Average	Discount_Fee

SCENARIO: Grouping on Institute and Mode, and finding Count of students, Sum and Average of Discountfee

Group By

Specify the columns to group by and one or more outputs.

☐ Basic ☒ Advanced

Group by

InstituteID

ModelID

Add grouping

New column name	Operation	Column
Cnt_Students	Count Rows	
Sum_Fee	Sum	Discount_Fee
Avg_Fee	Average	Discount_Fee

SCENARIO: Fill Down and Fill Up Practice

Have null values in the columns and specify Fill Up / Down so that the last value copied in the null cells up and down

Note: Go to the appended data of Courses_New and Courses_Upcoming

CourseID → Fill Down

Duration → Fill Up

Note: Highlight the column → right click

SCENARIO: Implementing Unpivot Operation

Refer to initial **Budget. Xlsx** file to see the unpivot real-time advantage

Scenario: General scenario

PARTY

PARTYID	SALCOMPONENT	SALAMOUNT
1	HRA	10000
1	DA	10000
1	TA	10000
2	HRA	20000
2	DA	20000
2	TA	20000

ROW TO COLUMN CONVERSION IS PIVOT,
REVERSE IS UNPIVOT--COLUMNS TO ROWS

PARTYID	HRA	DA	TA
1	10000	10000	10000
2	20000	20000	20000

a) Construct the above table

	Column1	Column2	Column3	Column4	*
1	partyid	hra	ta	da	
2	1	10000	20000	20000	
3	2	20000	10000	10000	
4	3	30000	20000	30000	

Name: PIVOT_DATA_TABLE

- b) Go to table → Set First ROW as headers
c) Choose columns (HRA /TA/ DA) → Rc → UnPivot columns
Now it shows column conversion to rows.

Which columns are suitable for Unpivot?

Columns which has measurable more measurable [recommended] values and you want to see as row values for better data calculations.

Ex: HRA, TA, DA

JAN, FEB, MARCH etc...

Address1, Address2, Address3 ...

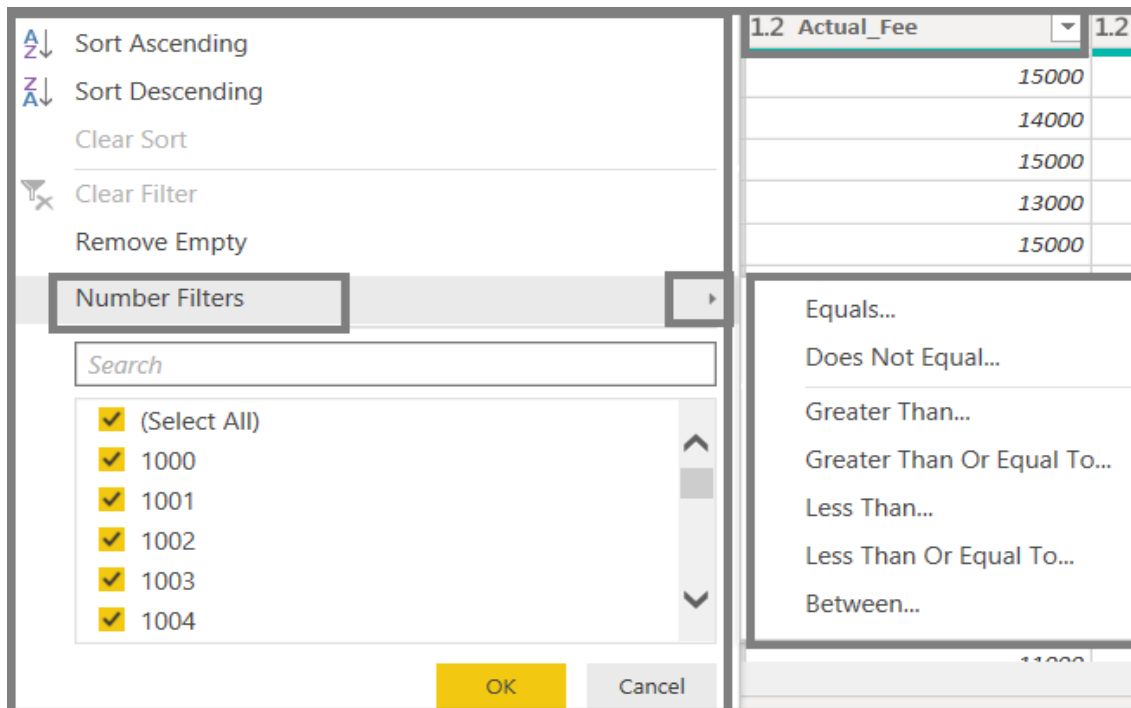
Vinay Tech House

d)Filter properties

There are three major types of filter properties. These help to have required data in the query.
Parameters and other intermediaries filtered here.

- a) Numeric filter properties
- b) String filter properties
- c) Date filter properties

Numeric Filters:



a)Sort Ascending: Keep the values in ascending order

b)Sort Descending: Keep the values in descending order

a)Clear Sort: Clearing the sort order specified

a)Clear Filter: Clearing the filters applied

a)Remove Empty: Empty values removal

Filters:

Keeping the filters on values. There are two types of filters.

- a) **Static filter:** Choosing required values by check / uncheck
- b) **Conditional filter:** Applying a condition to have required values.

Example: Keep only the actual fee values between 12000 and 14000

Number Filters → Between

Filter Rows

Apply one or more filter conditions to the rows in this table.

☒ Basic ☐ Advanced

Keep rows where 'Actual_Fee'

is greater than or equal to 1.2 12000

☒ And ☐ Or

is less than or equal to 1.2 14000

OK Cancel

STRING FILTERS

AB C ModelID AB C LocationID

Sort Ascending
Sort Descending
Clear Sort

Clear Filter
Remove Empty

Text Filters

Search

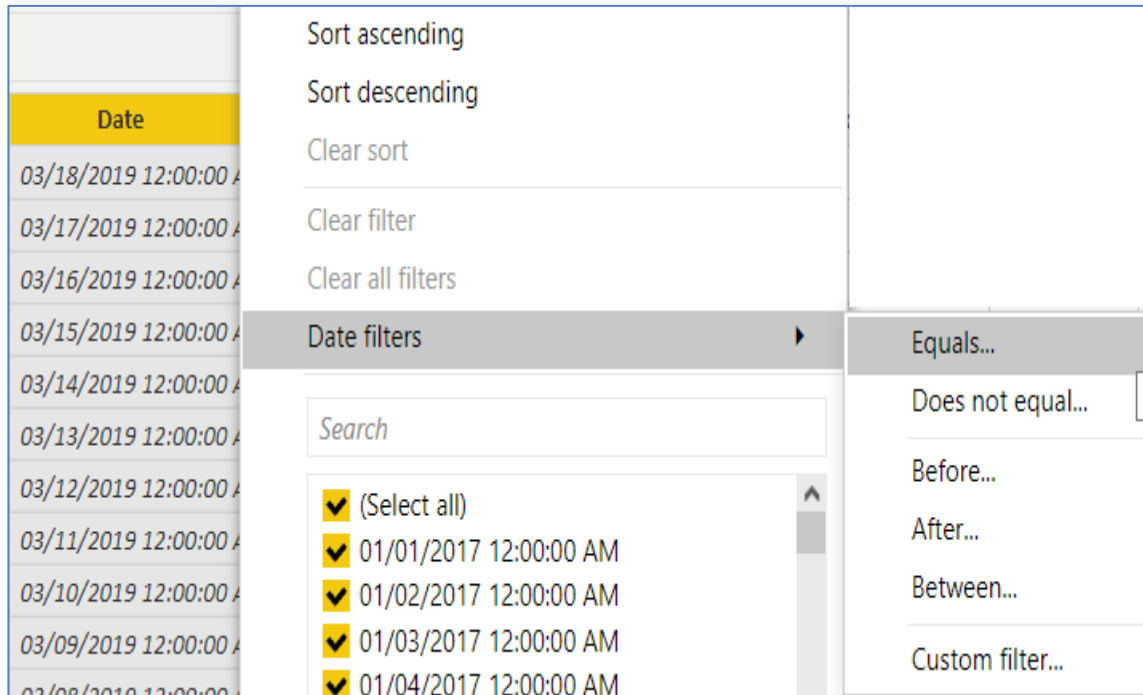
☒ (Select All)
☒ BZA
☒ HYD
☒ VIZAG

OK Cancel

Equals...
Does Not Equal...
Begins With...
Does Not Begin With...
Ends With...
Does Not End With...
Contains...
Does Not Contain...

Examples:

- Get the students whose name starts with 'v'
- Get the students whose names do not contain 'marry'

Date Filter:**Custom Filter:**

In case the above filters not directly supporting your operation, to construct own formula, we go for custom filtering.

Example: Display 2019 March 1 and 2019 March 19 data

Filter Rows

Apply one or more filter conditions to the rows in this table.

☒ Basic ☐ Advanced

Keep rows where 'Enquiry_Date'

equals

☐ And ☒ Or

equals

OK Cancel

PROJECT NUMBER -1 [GETTING, POWER QUERY SHAPING and MODELLING]

Greetings from VINAYTECH!!

Congratulations on making it to the next project of our Power BI Course Curriculum!!

PFA the Case Study.

Kindly complete the given Case Study and send it back within a Week.

Business Case:

Analyze Venture capital funding data for some of the popular companies and present insights and observations from the data. Please understand that you have to make sense of the limited data provided and call out any assumptions you make.

Data Provided:

The attached excel workbook has 2 sheets:

Funding Data: This sheet contains the details of the company being funded, date of funding, raised amount (\$) and the round series of the funding

State Map: Contains the City to state mapping

Deliverables Expected:

Integration of two sheets

Data cleansing, missing value treatment

What is the story behind the data? What are some of the findings and insights from data exploration (This can be in PowerBI environment or anything better for visual story telling?)

Kindly reach out in case of any queries or clarifications.

**Regards,
SRI VINAY TECH HOUSE.**